Change of basis in a linear space

Food for Thought Problem 1. Let V be a linear space without any additional structure on it (i.e., there is no norm, inner product, or the concept of orthogonality). Consider the linear operator $\mathsf{A} : V \to V$ acting on V. Let $\mathbf{f}_1, \ldots, \mathbf{f}_n$ be a basis in V. One can define the matrix elements of A in this basis as follows:

$$\mathsf{A}\mathbf{f}_j =: \sum_{i=1}^{n} a_{ij}\, \mathbf{f}_i \ ,$$

so that the matrix of A in this basis is $\underline{\underline{A}} = (a_{ij})$ (see equation (7) on page 457).

(a) Let $\mathbf{u} = \sum_{j=1}^{n} u_j\, \mathbf{f}_j$. Find the components of $\mathsf{A}\mathbf{u}$ in the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$; in other words, if $\mathbf{v} = \mathsf{A}\mathbf{u}$, find the numbers v_i such that $\mathbf{v} = \sum_{i=1}^{n} v_i\, \mathbf{f}_i$.

Solution:

$$\mathbf{v} = \mathsf{A}\mathbf{u} = \mathsf{A}\sum_j u_j\, \mathbf{f}_j = \sum_j u_j\, \mathsf{A}\mathbf{f}_j$$

$$= \sum_j u_j \sum_i a_{ij}\, \mathbf{f}_i = \sum_i \left(\sum_j a_{ij}\, u_j \right) \mathbf{f}_i \ .$$

Comparing this with $\mathbf{v} = \sum_i v_i\, \mathbf{f}_i$, we see that

$$v_i = \sum_j a_{ij}\, u_j \ .$$

Recalling the definition of a multiplication of two matrices, we can interpret this equality as follows: if the vectors $\mathbf{u}$ and $\mathbf{v}$ are written as column vectors (i.e., as $n\times 1$ matrices, where $n = \dim V$), and $\underline{\underline{A}} = (a_{ij})$ is an $n \times n$ matrix, then the column vector $\mathbf{v} = \mathsf{A}\mathbf{u}$ is the product of the matrix $\underline{\underline{A}}$ and the column vector $\mathbf{u}$.

(b) Let B be another linear operator with matrix elements (b_{ij}) in the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$. Find the matrix elements of the operator AB in the same basis. Here AB stands the composition of the operators in the sense that $(\mathsf{AB})\mathbf{u} := \mathsf{A}(\mathsf{B}\mathbf{u})$.

Solution:

$$(\mathsf{AB})\mathbf{u} = \mathsf{A}(\mathsf{B}\mathbf{u}) = \mathsf{A}\left(\mathsf{B}\sum_j u_j\, \mathbf{f}_j \right) = \mathsf{A}\left(\sum_j u_j\, \mathsf{B}\mathbf{f}_j \right) = \mathsf{A}\left(\sum_j u_j \sum_i b_{ij}\, \mathbf{f}_i \right)$$

$$= \mathsf{A}\left(\sum_{j,i} u_j\, b_{ij}\, \mathbf{f}_i \right) = \sum_{j,i} u_j\, b_{ij}\, \mathsf{A}\mathbf{f}_i = \sum_{j,i} u_j\, b_{ij} \sum_k a_{ki}\, \mathbf{f}_k = \sum_{j,i,k} u_j\, b_{ij}\, a_{ki}\, \mathbf{f}_k$$

$$= \sum_k \left[\sum_j \left(\sum_i a_{ki}\, b_{ij} \right) u_j \right] \mathbf{f}_k = \sum_k \left[\sum_j \left(\underline{\underline{A}}\, \underline{\underline{B}} \right)_{kj} u_j \right] \mathbf{f}_k \ ,$$

therefore the kth component of $(\mathsf{AB})\mathbf{u}$ is

$$[(\mathsf{AB})\mathbf{u}]_k = \sum_j \left(\underline{\underline{A}}\,\underline{\underline{B}}\right)_{kj} u_j \ .$$

This together with the result from part (a) allows us to conclude that the matrix of the linear operator AB is $\underline{\underline{A}}\,\underline{\underline{B}}$.

(c) Let $\mathsf{C} : V \to V$ be an *invertible* linear operator, i.e., a linear operator that has an inverse C^{-1}, so that $\mathsf{C}(\mathsf{C}^{-1}) = \mathsf{C}^{-1}\mathsf{C} = \mathsf{I}$ (where I is the identity operator: $\mathsf{I}\mathbf{u} = \mathbf{u}$ for any $\mathbf{u} \in V$, the matrix of I in any basis is the unit matrix $\underline{\underline{I}} = (\delta_{ij})$). Let the operators C and C^{-1} have matrices $\underline{\underline{C}} = (c_{ij})$ and $\underline{\underline{D}} = (d_{ij})$, in the basis $\mathbf{f}_1, \ldots, \mathbf{f}_n$:

$$\mathsf{C}\mathbf{f}_j = \sum_{i=1}^{n} c_{ij}\,\mathbf{f}_i \ , \qquad \mathsf{C}^{-1}\mathbf{f}_j = \sum_{i=1}^{n} d_{ij}\,\mathbf{f}_i \ .$$

Show that the matrix $\underline{\underline{D}}$ corresponding to C^{-1} is equal to the inverse of the matrix $\underline{\underline{C}}$ corresponding to C, i.e., $\underline{\underline{D}} = \underline{\underline{C}}^{-1}$.

<u>Solution</u>: By the result of part (b), the operator $\mathsf{C}\,(\mathsf{C}^{-1})$ has matrix $\underline{\underline{C}}\,\underline{\underline{D}}$, so $\mathsf{C}\,(\mathsf{C}^{-1}) = \mathsf{I}$ implies that $\underline{\underline{C}}\,\underline{\underline{D}} = \underline{\underline{I}}$, i.e., that $\underline{\underline{D}} = \underline{\underline{C}}^{-1}$.

(d) Let us use the operators C and C^{-1} to define a new basis in V. Let the vectors of the new basis $\widetilde{\mathbf{f}}_1, \ldots, \widetilde{\mathbf{f}}_n$ be defined by

$$\mathbf{f}_j = \mathsf{C}\,\widetilde{\mathbf{f}}_j = \sum_{i=1}^{n} c_{ij}\,\widetilde{\mathbf{f}}_i \ ;$$

clearly, this implies that

$$\widetilde{\mathbf{f}}_j = \mathsf{C}^{-1}\,\mathbf{f}_j = \sum_{i=1}^{n} d_{ij}\,\mathbf{f}_i \left(= \sum_{i=1}^{n} (\underline{\underline{C}}^{-1})_{ij}\,\mathbf{f}_i \right) \ .$$

Prove that if $\widetilde{u}_j$ are the components of $\mathbf{u}$ in the new basis, i.e., $\mathbf{u} = \sum_{j=1}^{n} \widetilde{u}_j \widetilde{\mathbf{f}}_j$, then $\widetilde{u}_i = \sum_{j=1}^{n} c_{ij} u_j$.

<u>Solution</u>:

$$\mathbf{u} = \sum_j u_j\,\mathbf{f}_j = \sum_j u_j\,\mathsf{C}\,\widetilde{\mathbf{f}}_j = \sum_j u_j \sum_i c_{ij}\,\widetilde{\mathbf{f}}_i = \sum_i \left(\sum_j c_{ij} u_j \right) \widetilde{\mathbf{f}}_i$$

– compare with

$$\mathbf{u} = \sum_j \widetilde{u}_j\,\widetilde{\mathbf{f}}_j$$

to get

$$\widetilde{u}_i = \sum_j c_{ij}\,u_j \ .$$

(e) Now we want to find the matrix of the operator A in the new basis $\widetilde{\mathbf{f}}_1, \ldots, \widetilde{\mathbf{f}}_n$. Define the matrix elements $\widetilde{a}_{ij}$ of A in the new basis by

$$\mathsf{A}\widetilde{\mathbf{f}}_j =: \sum_{i=1}^{n} \widetilde{a}_{ij}\, \widetilde{\mathbf{f}}_i \ ,$$

and let $\underline{\underline{\widetilde{A}}} = (\widetilde{a}_{ij})$. Show that $\underline{\underline{\widetilde{A}}} = \underline{\underline{C}}\,\underline{\underline{A}}\,\underline{\underline{C}}^{-1}$.

<u>Solution</u>: Using the formulas from part (d) that connect the bases $\mathbf{f}_i$ and $\widetilde{\mathbf{f}}_i$, we obtain

$$\begin{aligned} \mathsf{A}\,\widetilde{\mathbf{f}}_j &= \mathsf{A} \sum_i d_{ij}\, \widetilde{\mathbf{f}}_i = \sum_i d_{ij}\, \mathsf{A}\, \widetilde{\mathbf{f}}_i \\ &= \sum_{i,k} d_{ij}\, a_{ki}\, \widetilde{\mathbf{f}}_k = \sum_{i,k,l} d_{ij}\, a_{ki}\, c_{lk}\, \widetilde{\mathbf{f}}_l = \sum_l \left(\sum_{i,k} c_{lk}\, a_{ki}\, d_{ij} \right) \widetilde{\mathbf{f}}_l \ . \end{aligned}$$

Comparing this with $\mathsf{A}\widetilde{\mathbf{f}}_j = \sum_l \widetilde{a}_{lj}\, \widetilde{\mathbf{f}}_l$, we see that

$$\widetilde{a}_{lj} = \sum_{i,k} c_{lk}\, a_{kl}\, d_{ij} \ ,$$

i.e., that $\underline{\underline{\widetilde{A}}} = \underline{\underline{C}}\,\underline{\underline{A}}\,\underline{\underline{C}}^{-1}$ (recall that $\underline{\underline{C}}^{-1} = (d_{ij})$).